Abstract

A feasibility patient preference quasi-experimental study was conducted of group cognitive analytic music therapy (G-CAMT) with mentally disordered offenders. Participants either chose or were randomised to 16 sessions of manualised G-CAMT (N=10) plus treatment as usual (TAU) or TAU alone (N=10). Self-rated and staff-rated outcomes were assessed at baseline, post-intervention and 8-weeks post-intervention. Residency was assessed at 2-year follow-up. Results indicate that G-CAMT was easily implemented; 9/10 participants completed G-CAMT and attendees had high satisfaction with the approach. Session attendance was high; 4/10 participants attended all sessions. At the 8-week follow-up, 3/9 G-CAMT participants had reliable reductions (i.e. statistically reliable pre to 8-week follow-up change results) in intrusive/possessive behaviours and fear of separation/abandonment. On the staff-rated outcome measure G-CAMT participants as a group were statistically significantly friendlier compared to TAU at 8-week follow-up (U = 0.50, p = 0.009, d = 1.92, CI 0.44 to 3.11). There were no differences between the arms in terms of residency outcomes at 2-year follow-up. The study is discussed in terms of G-CAMT's theoretical grounding and high acceptability. The study is limited by its small sample size, but indicates the possibility of progressing onto a full trial.

Keywords: cognitive analytic therapy, patient preference; high secure hospital; forensic music therapy

Background

Duggan et al. (2006) emphasised the importance of providing acceptable and evidence-based treatments for incarcerated seriously mentally disordered offenders (MDOs). An MDO is a person who has a disability or disorder of the mind and has been found guilty of committing a criminal offence (Fassaert et al, 2016). MDOs are detained in secure hospitals for treatment until rigorous risk assessment procedures determine that they are sufficiently recovered to move to conditions of lesser security - the average stay being over six-years in high secure hospitals (Butwell, Jamieson, Leese & Taylor, 2000). The treatment needs of MDOs markedly differ from both patients in generic mental health services and prisoners in correctional settings (Thomas, 2005). This is because MDOs frequently have personality disorder (PD) and co-morbid severe mental illness; McGauley and Bartlett (2009) noted that MDOs have been historically labelled 'too ill/disturbed' for psychotherapy. Whilst in recent years, psychotherapies have been increasingly integrated into MDO multi-disciplinary treatment pathways; forensic psychotherapy still remains a relatively new approach (Gilligan, 2015).

Evidence base for music therapy with schizophrenia and with offenders

The evidence base for music therapy (MT) as an adjunct treatment for schizophrenia has graduated from initial qualitative case studies (e.g. psychodynamic/analytic MT; Glyn, 2009) to randomized and controlled clinical trials (RCTs) and associated meta-analyses. For example, Talwar et al. (2006) randomized (n=81) community-based patients with acute psychosis to either 12-sessions of individual MT (plus treatment as usual; TAU) or TAU. Attendance was high (median MT attendance; eight sessions) with a trend towards improved symptom scores for MT, especially regarding general symptoms of schizophrenia. Tseng et al. (2016) conducted an updated meta-analysis of the role of MT as an adjunct schizophrenia

treatment and investigated effective clinical characteristics. Outcomes were significantly improved for those receiving MT for negative and positive symptoms, regardless of total treatment duration or number/frequency of sessions. Chen, Leith, Aarø, Manger and Gold (2016) completed a meta-analysis of MT for offenders in correctional settings. MT (regardless of theoretical orientation) was effective in promoting self-esteem (Hedges' g = 0.55) and social functioning (Hedges' g = 0.35), with anxiety and depression outcomes dependent on length of treatment. Overall, whilst there is an evidence base for the use of MT as an adjunct treatment in schizophrenia, there is a lack of sufficient rigorously controlled outcome research of MT with MDOs.

Cognitive analytic therapy; theoretical framework

Cognitive analytic therapy (CAT) uses a time-limited, collaborative and relational therapeutic approach to facilitating change (Ryle, Kellett, Hepple & Calvert, 2014), whose theory and clinical procedures and practice are well established (Ryle & Kellett, 2018). Theoretically, CAT integrates personal construct and object relations theory (Ryle, 1985), to assert that mental representations of the self and others are founded in (and then maintained) by developmental early interactions with care providers and other significant figures (Ryle & Kerr, 2002). CAT theory consists of three theoretical foundation stones; *reciprocal roles, target problem procedures* and the *multiple self-states model* (MSSM). Internalized early object relations are termed '*reciprocal roles*' and these influence the manner in which an MDO might predict, establish, experience, maintain and respond to relationships across the lifespan. Reciprocal roles can range from self to self (i.e. the relationship the MDO has with themselves), self to other (i.e. how the MDO relates to others, including the music therapist) and other to self (i.e. what the MDO elicits from others and also how they experience others, including the music therapist). Reciprocal roles therefore represent the analytic component

of the CAT model (Ryle & Kellett, 2018). CAT conceptualizes the crimes committed by MDOs as a consequence of childhood neglect, abandonment and abuse, leading to the internalization and then acting out of pathological reciprocal roles. Green, Brown & Chew's (2017) meta-analysis ($N = 2215$) showed that individuals with histories of childhood maltreatment and current psychosis had double the risk of perpetrating violence. In CAT, *target problem procedures* (*TPPs;* Ryle & Kerr, 2002) are labelled as traps (i.e. vicious circles), snags (i.e. self-sabotage) and dilemmas (i.e. either or dilemmas). TPPs are used to summarise the procedural sequences that explain the manner in which the MDO currently thinks, feels and behaves, and in so doing highlights the present day consequences of these actions. Procedures represent the cognitive component of CAT (Ryle & Kellett, 2018).

The procedural sequence object relations model (PSORM) was an integration of the cognitive and analytic aspects of the CAT model that highlighted that procedures were most often the product of reciprocal role activation (Ryle, 1991). In CAT, narrative and sequential diagrammatic reformulations are collaboratively produced to enable better recognition of historical antecedents of roles occupied in relationships and how procedures often maintain and link the roles together in the here and now (Ryle & Kellett, 2018). The multiple self-states model (MSSM; Ryle 1997) was a further development of the PSORM, which attempted to account for the identity disturbance and rapid switching between extreme states that frequently presented in complex forensic cases (Pollock et al. 2001). The MSSM implies a polyphonic self, emphasising how childhood trauma can create distinct and marked separation between varieties of dysfunctional key states, which are maintained through ongoing dissociation (Ryle, 2007). For example, Quimby and Putnam (1991) found that trait dissociation was positively correlated with sexual aggression among adult psychiatric inpatients.

The three key theoretical aspects of CAT therefore offer potential utility in

conceptualizing the state-shifting (e.g. chronic placation suddenly switching into a raging physical attack), reciprocity (e.g. occupying both abusing and abused roles) and procedural elements (e.g. abusive actions having interpersonal consequences, such as eliciting rejection from social groups) of the past crimes and current relational styles of MDOs that maintain ongoing distress and risk. The analytic nature of the CAT model also enables the opportunity to analyse when 'role enactments' in the therapeutic relationship with the therapist mirror both past relationships and current relationship patterns with staff (Ryle & Kellett, 2018). Clinically, CAT uses a three-phase approach to change; (1) a *reformulation* stage during which the patient's presenting problem are reformulated via narrative and sequential diagrammatic reformulations, (2) a *recognition* stage to enable the patient to better recognise when they are caught in procedures or roles and (3) the final stage of *revision* in which the patient and therapist work collaboratively to explore how to respond differently by changing roles and procedures. The change methods of CAT used during the revision phase are varied and comprehensive and can be taken from other psychotherapeutic modalities (Ryle & Kellett, 2018). What is essential is that any change method (or 'exit' in the language of CAT) is grounded in the sequential diagrammatic reformulation. This ensures that exits are patient centred and also within the patient's individual capacity for self-development, known as the zone of proximal development (ZPD; Vygotsky, 1978). Knabb, Welsh and Graham-Howard (2011) therefore highlighted the strengths of CAT in a forensic setting as being the model's strong focus on conceptualisation, an ability to understand MDO intra-psychic processes and analysis of offender-victim dynamics.

Current study; background, choice of measures and aims

This paper reports the results from a feasibility patient preference quasi-experimental research study of a manualised MT with MDOs delivered in groups, theoretically grounded in cognitive analytic therapy (Ryle & Kerr, 2002) - thereafter named group cognitive analytic

music therapy (G-CAMT; Compton Dickinson, 2015). Manualization enables well-defined MT practice frameworks to be tested during controlled studies (Carr, Odell-Millar & Priebe, 2013). Whilst some studies have examined the effectiveness of CAT in forensic populations (e.g. Pollock & Belshaw, 1998), such studies have lacked sufficient methodological control. As CAT offers a framework for better understanding the origins and current maintainers impairing effective relating to self and others (Ryle & Kerr, 2002), the primary psychometric outcome measure used in the present study was therefore one of relatedness (i.e. the Person's Relatedness to Others Questionnaire-2 (PROQ-2; Birtchnell & Evans, 2004). The PROQ2 has been used previously as a primary outcome measure in a CAT study (Birtchnell, Denman & Okhai, 2009), as well as in a forensic treatment setting (Birtchnell et al. 2009). To increase the study's internal validity, this was supplemented by an observer-rated measure designed for use in secure treatment settings called the chart of interpersonal reactions in a closed living environments (CIRCLE; Blackburn & Glasgow, 2006). There is evidence of CAT reducing pathological dissociation via personality integration (Kellett, 2005) and so a secondary measure of dissociation was also used.

As a feasibility trial (Bowen et al. 2010), this study sought to answer the following questions; was G-CAMT well tolerated as an intervention (i.e. the percentage of treatment sessions attended and the overall dropout rate), easily implemented (i.e. how much support the music therapists required to implement treatment), popular (i.e. the rate of patients choosing G-CAMT) and a satisfactory and acceptable treatment when included within multi-disciplinary treatment for MDOs in a high secure hospital (as measured by responses to a patient satisfaction survey). The study also sought to complete some limited-efficacy testing (Bowen et al. 2010) of G-CAMT (i.e. within and between group effect size calculations, but with limited statistical power), define the rate of change at an individual level for a patient going through G-CAMT and to also assess the long-term residency of participants. The value

and contribution of feasibility trials prior to conducting main studies is widely acknowledged (Lancaster et al. 2010) and this study sought to test the feasibility of a widely used (but under researched) forensic model of MT in a high secure setting (Pollock, Stowell-Smith & Gopfert, 2006).

When it is impossible to mask patients to treatment allocation (as in psychotherapy trials), a potential confound is created of recording treatment effects more representative of patient's preferences rather than true therapeutic efficacy (Howard & Thornicroft, 2007). In order to maximise treatment choice, the methodology used in the current study was therefore a patient preference methodology (Torgerson & Sibbald, 1998). Whilst patient preference for use of music is usually a consideration in MT outcome studies (Kamioka, et al. 2014), only one previous MT study has attempted this methodology (Gold et al. 2014), and this highlighted the recruitment challenges for implementing research in prisoner populations. The primary aim for this study was to consider whether G-CAMT was an easily implemented and satisfactory intervention. The secondary aims were (a) to compare the dropout rate for G-CAMT with the CAT dropout rates reported n the literature, (b) test whether G-CAMT could differentially improve relational and empathic abilities and differentially reduce dissociative symptoms and (d) to assess long-term residency outcomes.

Method

Design

Ethical and governance approval was achieved (ref: East Midlands; Nottingham1: 1/HO0403/4). The G-CAMT intervention was delivered between September and December 2011. The methodology was a patient preference trial (Torgerson & Sibbald, 1998), with randomization conducted independently via a computer generated randomization process. Partial randomization was based on the following: (a) patients with no strong preference

consenting to randomization, (b) patients with a preference for either G-CAMT or control, but who still consented to randomization and (c) patients who chose their preference for either G-CAMT or TAU and therefore refused randomization. Prior to attending G-CAMT, participants attended an initial individual session to explain the approach and to define target problems and treatment goals (Ryle & Kerr, 2002). Multi-disciplinary TAU consisted of medication, plus recreational music-as-listening groups, occupational and speech therapy, offence related psychology programmes and physical activities such as gym, swimming and horticulture.

Participants and allocation

The study sample were male patients (aged 18-65) who had been convicted of a violent offence and deemed to have serious mental disorder and so were detained under the Mental Health Act (2008) in a high secure hospital. All participants had a diagnosis of schizophrenia or serious mental disorder, with or without comorbid PD. Recruitment took place across five inpatient male wards. Inclusion criteria were an IQ of > 70, hospital residency for over 6-months, sufficient comprehension to complete outcome measures, psychotic illness not precluding the completion of outcomes measures and some initial interest in using MT. Exclusion criteria were a diagnosis of organic brain disease, serious physical illness, IQ < 70, physical disability impairing movement to a degree which precluded MT participation, already attending music/art therapy, currently actively and floridly psychotic and admitted within the past 6-months.

The CONSORT summary (Moher et al. 2010) is in Figure 1. From a potential pool of N=102, then n=35 did not meet inclusion criteria, n=15 actively declined to participate and n=32 provided other reasons for not participating (e.g. the MDOs informed the research team of their lack of availability, due to attending other activities and existing treatment options).

Twenty patients consented to participate and four chose to be randomized. There were 4 dropouts before commencement of the trial (1 randomized to TAU and 3 that had expressed a preference for TAU), 2 gave no reason and 2 became actively psychotic. The research sample (n=16) therefore comprised 3 participants that chose to be randomized and 13 that expressed a treatment preference (9 chose G-CAMT and 4 chose TAU). When randomized and preference patients were combined (see analysis section), 10 received G-CAMT and 6 received TAU. Patients allocated to G-CAMT were then further randomly allocated to participate in one of two groups (5 participants started in each group).

Music therapy

G-CAMT was designed for usage with MDOs with primary diagnoses of schizophrenia and/or comorbid personality disorder. The three-phase iterative approach to G-CAMT treatment manual development was consistent with Medical Research Council guidelines for the development of complex interventions (Campbell et al. 2000; 2007). Phase one involved an initial 2-year modelling phase to develop the group-work model from a series of individual CAMT cases, as cited in the National Institute of Health and Care Excellence consultation for antisocial PD (NICE, 2007). Phase two involved a G-CAMT pilot group (12-sessions), with participants acting as their own controls (Lawday & Compton, 2013). This enabled a compatibility check of G-CAMT with TAU in the hospital [ward-based multi-disciplinary treatment (MDT) programmes] environment. The third phase involved acting on the feedback from phase 2, and then more formal testing utilizing the current feasibility quasi-experimental study. G-CAMT in the current study therefore consisted of a 16-session manualised intervention with each session of 90-minutes duration. G-CAMT has four structured stages each building on the previous one to develop manageable layers of learning; (1) establish a safe environment for group work, with specific musical mindfulness, harmonic and grounding techniques (2) emotional recognition: developing self-awareness and

understanding of reciprocal roles, (3) distress tolerance work, introducing more musical resources and chromatic dissonance that represent enhanced ability to safely address interpersonal conflict (4) tolerating endings with reduced impulsivity, experiencing mourning and saying 'farewell' (Compton Dickinson & Hakvoort 2017). These stages are grounded on the reformulation, recognition and revision structure of CAT (Ryle & Kellett, 2018) and guidance is provided in using the relevant the CAT tools. Table 1 provides a stage and session-by-session description of the intervention and its fidelity to CAT theory. In order to deliver G-CAMT, two post-graduate qualified health and care profession council (HCPC) registered music therapists attended an introductory two-day training in the CAT model. They then received weekly clinical supervision (with a registered music therapist and accredited cognitive analytic psychotherapist and supervisor) to ensure adherence to the G-CAMT treatment manual. Each group (n=5) was run by a single music therapist. The MT supervisor received 'supervision of supervision' by an accredited CAT supervisor as an additional quality assurance process.

The main active ingredient of G-CAMT across the four stages is jointly-created musical improvisation - this concept is drawn from activity theory (Vygotsky, 1978), which is a central tenet of CAT (Ryle & Kerr, 2002). Therefore, G-CAMT is based on a dialogical approach (Bakhtin, 1981) to spontaneous music creation and innovation. Creative self-expression is facilitated during G-CAMT via musical and verbal dialogue that develops between the music therapist and patients. Risks of violence are managed through the treatment stages by gradually increasing the musical resources and techniques in a 'scaffolded approach' to enable positive changes in relational abilities (Ryle & Kerr, 2002). Within this manualised model specific novel techniques (termed the sound print and safety call) were developed to promote positive relating to both the creative medium, facilitators and other group members. Previous qualitative analysis of semi-structured interviews conducted

with the music therapists has revealed that the use of the manual developed group cohesiveness and enabled analysis and safe management of 'risky' behavioural enactments, rupture-repair sequences and improved relational awareness (Compton Dickinson, 2015).

Outcomes

Psychometric outcomes were collected for G-CAMT participants at baseline (prior to self-preference or random allocation), then after the termination of the 16-session group intervention and again at 8 weeks post-intervention (24 weeks from baseline). Patients allocated to TAU had the measures taken at matched corresponding time intervals. Hospital residency was assessed at 2-years follow-up. G-CAMT also completed a measure of service satisfaction. All self-report outcome measures were administered and collected independently by a research assistant masked to allocation. The observational measure was completed by ward nursing staff.

The Persons Related to Others Questionnaire-2 (PROQ-2; Birtchnell & Evans, 2004). This scale measures eight domains of relatedness. There are four options for each of the 96 items providing a score range of 0-3 (rarely true, sometimes true, quite often true and nearly always true). For each octant scale, the score range is 0-30. The total PROQ-2 score has a maximum score of 240, and a reduction to the total score would be indicative of positive relational change. The PROQ-2 subscales conceptually correspond to the octants of an octagon constructed around a horizontal (close/distant) axis, a vertical (upper/lower rank) one, and four intermediate positions. (Figure 2). The four classes on the PROQ-2 are termed the upper close (UC, $\alpha = 0.85$), lower close (LC, $\alpha = 0.85$), upper distant (UD, $\alpha = 0.73$) and lower distant (LD, $\alpha = 0.84$). Each of the eight positions of the interpersonal octagon has a two-word name, the first word referring to one or other of the polarities of the vertical axis, (Upper and Lower) and the second referring to the horizontal one (Close and Distant). For the

four primary domains, the word neutral is used, hence they are termed neutral close (NC, $\alpha =$ 0.82), neutral distant (ND, $\alpha = 0.79$), upper neutral (UN, $\alpha = 0.75$) and lower neutral (LN, $\alpha =$ 0.80). The PROQ-2 has been subject to reliability and validity testing with depressed patients (Birtchnell, Falkowski & Steffert, 1992) and forensic participants (Birtchnell & Shine, 2000). In the depressed sample, the mean total negative score was significantly higher for participants diagnosed with depression, as were the mean negative scores for lowerness (p <0.001), closeness (p< 0.005) and distance (p<0.001). In the forensic sample, 29% of prisoners scored greater than 130 on the full-scale, item-total correlations ranged from .3 to .7 and test-retest reliabilities across the sub-scales ranged from .63 to .93 and .90 for the full-scale. PROQ-2 total mean scores are 98.5 ($SD = 26.9$) for non-patients and 132.7 ($SD =$ 23.9) for patients (Birtchnell & Evans, 2004), with correlations generally >.4 with a personality disorder diagnosis (Birtchnell & Shine, 2000).

Basic Empathy Scale (BES; Jolliffe & Farringdon, 2006a).

This 20-item scale measures affective ($\alpha = 0.84$) and cognitive empathy ($\alpha = 0.71$) abilities and the two-factor solution has been tested with confirmatory factor analysis (Joliffe & Farringdon, 2006a). Each BES item asks participants to respond on a Likert scale from 1 (strongly disagree) to 5 (strongly agree), so that sum scores can range from 20-100. The test-retest over a 7-week period for the affective empathy was r = .79 and for cognitive empathy r = .61.   Males have a mean BES score of 64.3 ($SD = 9.8$) and females a mean BES score of 75.3 (SD = 8.3). The size of the BES difference between females and males in affective empathy ($d = 1.33$) is much larger than for cognitive empathy ($d = 0.66$; Joliffe and Farringdon, 2006b). An increase in the overall mean for the BES would be indicative of an increase in the ability to empathise with others.

Multi-Scale Dissociation Inventory (MDI; Briere, 2002).

This 30-item scale measures six domains of dissociation: disengagement ($\alpha = 0.83$), de-personalization ($\alpha = 0.90$), de-realization ($\alpha = 0.91$), emotional constriction ($\alpha = 0.94$), memory disturbance ($\alpha = 0.74$) and identity dissociation ($\alpha = 0.75$). Each item is rated according to its frequency of occurrence over the prior month, using a scale ranging from 1 (never) to 5 (very often), and so summary scores range from 30-150. Reductions MDI scores would imply a reduction to dissociative symptomatology. Exploratory factor analysis has validated the subscale structure and demonstrated that clinical trauma samples score significantly higher on the MDI and that the subscales predict prior trauma exposure with significant R^2 values ranging from 0.7 -. 31. (Briere, Weathers & Runtz, 2005). A raw identity dissociation subscale score of 15 or higher on the MDI successfully identified 93% of those with diagnosis of Dissociative Identity Disorder (Briere, 2002).

Chart of Interpersonal Reactions in Closed Living Environments (CIRCLE; Blackburn & Glasgow, 2006).

This is an observer measure of the interpersonal/social behaviours of hospitalised psychiatric inpatients (Blackburn & Renwick, 1996). The CIRCLE consists of 49 items that describe a variety of verbal and nonverbal social behaviours and each item is rated on a four point Likert scale from 0 (not at all) to 3 (usually or frequently) by two raters therefore creating a possible score range of 0-294 (i.e. 49x3x2). A high scale score represents the extent to which each interpersonal style (e.g. dominance, hostility and coercion) is observed by staff (Blackburn & Renwick, 1996) and so reductions in the CIRCLE total score would imply an improvement to the participants social functioning on the ward. The CIRCLE generates eight subscales (dominant, coercive, hostile, withdrawn, submissive, competitive, friendly and sociable) and was originally psychometrically validated in male forensic psychiatric inpatients (Blackburn, 1992). Vernham, Tapp and Moore (2015) found that the CIRCLE had good predictive validity for incidents of verbal/physical aggression and self-harm in high-secure forensic

inpatients at 12, 24 and 48 months; with the CIRCLE subscales significantly correlating between .09-.23 with incidents and area under the curve values for violent incidents at 12-months ranging from .54 to .78 for all CIRCLE subscales. Blackburn & Renwick (1996) illustrated that the eight subscales met circumplex geometric requirements in two samples of hospitalized male forensic psychiatric patients. Cronbach's alpha coefficients for the CIRCLE sub-scales ranges from .70-.92 (mean α =.79) and intra-class correlations range from .50-.67 (mean = .57; Vernham, Tapp and Moore, 2015). For the purposes of the current study, only four CIRCLE subscales were rated by a single member of staff: hostile, withdrawn, friendly and sociable.

Client Satisfaction Scale-8 (CSQ-8; Attkisson and Greenfield, 1994)

This is a valid and reliable measure of patient satisfaction with healthcare interventions. CSQ-8 scores range from 8-32, with higher scores indexing higher satisfaction. The CSQ-8 has good psychometric foundations, including high levels of internal consistency ($\alpha = 0.92$); good construct validity (e.g. patients dropping out of therapy within the first month were less satisfied than treatment completers r=0.37, p<0.01) and functions well in comparison to other service satisfaction measures. The CSQ-8 has been tested with confirmatory factor analysis to reveal a single factor of service satisfaction (Kelly et al. 2017). Twelve studies (N=8000) form the basis of the CSQ-8's development and evaluation (Rosenblatt & Attkisson, 1993).

Statistical analyses

Mann-Whitney U tests evaluated differences between the participants that stated a preference and those that were happy to be randomized. Because differences between preference and randomized participants were non-significant, we collapsed participants into the arms of the study. Mann-Whitney U tests then again used to compare the 'collapsed' study arms at pre-treatment on continuous and ordinal variables, with descriptive statistics comparing

categorical variables. In order to provide a conservative estimate of treatment effects, intention-to-treat analyses were conducted using the last observation carried forward method (Carpenter & Kenward, 2008). Descriptive statistics in outcome measures are provided for G-CAMT and TAU arms at each time point. Outcome analyses were then conducted at both individual and group levels. At the individual level, the size of psychological change achieved on PROQ-2 domains was categorised using the reliable change index (RCI; Jacobson & Truax, 1991) at termination and at follow-up. In accordance with recommendations (Evans et al. 1998), reliable improvement was recorded when an individual participant's pre-post score PROQ-2 sub-scale improved by equal to or more than 1.96 times the SEdiff (Connell et al., 2007). At the group level, within-group uncontrolled effect sizes (Cohen's d) were calculated to demonstrate the size of the effect in each arm and were calculated using the baseline-termination change score divided by the baseline standard deviation (Westbrook & Kirk, 2005). Effect sizes were reversed as necessary, so that a positive effect size always represented a favourable change, and classified effect sizes of 0.20 as "small", 0.50 as "medium", and 0.80 as "large" according to Cohen's guidelines (1992). Where the effect size reported is a minus score then therefore that would index a negative effect. Mann Whitney U tests were then conducted to assess the statistical significance of the change in outcomes scores observed post-treatment and at follow-up between the arms; the Bonferroni correction was applied given multiple testing and the significance cut-off was set at .025 (.05/2). At 2-year follow-up, residency (i.e. high vs. low security) was assessed. Individual participant data are available as supplementary online material.

Results

Results are presented in three sections; (a) description of the sample and comparisons of the study arms at baseline, (b) acceptability, attendance and satisfaction with G-CAMT and (c) comparison of G-CAMT and TAU on secondary outcomes on an individual and group basis.

Sample description, randomization checks and baseline comparisons

Table 2 displays the demographic and clinical characteristics of participants in the two arms

for whom baseline measures were available. Of participants choosing to be randomized

(N=4), all were White British with a diagnosis of schizophrenia and a history of substance

misuse, but no history of suicide attempts. For two of the randomized participants, the index

offence was manslaughter under grounds of diminished responsibility. Amongst participants

stating a preference (9/13 White British), 8/13 had a diagnosis of schizophrenia, 10/13 had a

history of substance misuse, 9/13 had attempted suicide and 3/13 had an index offence, which

involved a fatal killing. There were no significant differences between participants stating a

preference and those randomized in terms of age ($U = 18.00$, $p = .84$); years offending ($U =$

14.50, $p =.50. 50$); age first convicted ($U = 19.50$, $p = 1.0$), length of incarceration ($U =$

18.00, $p = .84$), PROQ-2 ($U = 16.50$, $p = .83$), MDI ($U = 15.50$, $p = .59$), BES ($U = 15.00$, p

$= .54$), CIRCLE hostile ($U = 3.50$, $p = .16$), CIRCLE withdrawn ($U = 8.00$, $p =.67$), CIRCLE

friendly ($U = 5.00$, $p =.28$) and CIRCLE sociable ($U = 7.50$, $p = .58$). As there were no

major significant differences apparent between baseline characteristics of participants that

were randomized and those with a treatment preference, all randomized and preference

participants were then collapsed into the respective arms of the study arms: G-CAMT or

TAU.

The subsequent baseline comparisons of collapsed G-CAMT vs. TAU participants

showed no significant between-group differences in terms of age at first conviction ($U =$

26.50, $p = 0.96$), duration of offending ($U = 28.50$, $p = 0.88$) or length of incarceration ($U =$

19.50, $p = 0.38$). TAU participants were, however, significantly more hostile at baseline ($U =$

2.00, $p = .006$). All G-CAMT participants (N=10) had a history of substance misuse, with

7/10 having a history of suicide attempts. There were no differences found between G-

CAMT and TAU in terms of the standard care provided from baseline to follow-up:

psychological input hours ($U = 13.50$, $p = .11$), physical activity hours ($U = 11.5$, $p = .07$),

recreational music hours ($U = 26.00$, $p = .96$) and other activities in hours ($U = 14.5$, $p = .15$).

Primary outcomes; acceptability, attendance and satisfaction with G-CAMT

A single patient allocated to G-CAMT attended one session, but was too psychotic to

continue with the intervention and dropped out. Therefore 1/10 participants dropped out of G-

CAMT after the groups had started. The dropout rate for CAT with complex clients in adult

mental health settings is 10-12% (Evans et al. 2016; Kellett et al. 2013). In terms of sessional

attendance at G-CAMT, the mean session attendance was 12.5 (*SD* = 5.23). Four out of ten

G-CAMT participants attended full treatment (i.e. 16 sessions), with a median session

attendance of 14 and a mode of 16. Mean CSQ-8 satisfaction scores were: satisfaction (3.51),

quality, (3.57), service (2.86), meeting needs (2.50), would recommend the service (3.33),

right amount of help (3.17), helpful dealing with the problem (3.00) and would return again

(3.00).

Secondary outcomes; psychometric scores and residency assessment

Table 3 displays individual participant RCI rates on the PROQ-2 domains. No single G-

CAMT participant reliably improved on baseline-termination comparisons, whereas two

TAU participants made reliable improvements in being less dominating (UN domain),

suspicious (ND domain) and controlling (UD domain). By 8-week follow-up however, two

G-CAMT participants had made reliable reductions in terms of being intrusive, restrictive or

possessive (UC domain) and in terms of their fear of separation and being alone (NC

domain). For those G-CAMT participants for whom reliable change on UC or NC domains

was statistically possible (given their baseline scores), 3/8 made reliable improvement on

either one or both of these domains by follow-up. In terms of the full-scale PROQ-2 reliable

change analyses, in the TAU group a single participant deteriorated at end of treatment and two participants made a reliable improvement by follow-up. In the G-CAMT group on the full-scale PROQ-2 reliable change analyses, no participants had reliably improved by end of treatment, one participant had reliably improved by follow-up and no participants reliably deteriorated either time point.

Table 4 reports the group level PROQ-2 outcomes over time and Figure 2 provides a visual representation of baseline-termination PROQ-2 group level outcomes using the interpersonal octagon. The effect size results in Table 4 demonstrate a small-to-medium positive treatment effect ($d = 0.47$) on the LN domain (reductions to shunning responsibility, self-denigrating behaviours) immediately following treatment and at 8-week follow-up for G-CAMT participants. At follow-up on the ND domain (i.e. suspiciousness), a medium effect size in the direction of deterioration for G-CAMT participants was found ($d = -0.48$), versus a medium effect size in the direction of improvement for TAU participants ($d = 0.67$). The full-score PROQ-2 outcomes demonstrate a small positive effect size at follow-up for G-CAMT participants ($d = 0.38$), whilst TAU participants demonstrated a minimal effect size ($d = 0.17$). There were no significant differences on full-score PROQ-2 scores between G-CAMT and TAU at end of treatment or at follow-up. An exception to this general trend was a significant difference ($p<.025$) between the arms on the ND domain (suspicious, uncommunicative and self-reliant behaviours; termination, $U = 7.00$, $p = .01$; 8-week follow-up, $U = 6.00$, $p = .007$).

As shown in Table 5, ability to empathise (BES scores) and tendencies to dissociate (MDI) did not significantly change over time in either of the treatment arms. In terms of staff ratings of study participants, the effect sizes were small (friendliness $d = -0.47$) to medium (sociability $d = -0.51$; hostility $d = 0.69$) on baseline to follow-up comparisons for G-CAMT participants. Comparisons of ward staff ratings of friendliness at follow-up also identified a

difference between the arms. The reductions in friendliness observed in TAU participants did not occur in G-CAMT participants over the same time period ($U = 0.50$, $p = .009$). In terms of residency, at 2-years post-follow-up 7/10 G-CAMT ($M = 19$ months, $SD = 11.52$) and 4/6 TAU participants ($M = 25.5$ months, SD = 4.65) had been moved to lower levels of secure hospital treatment.

Discussion

This study has reported on the first quasi-experimental study of MT in a secure hospital and therefore satisfies calls for controlled outcome studies of MT to be completed in such contexts (Duggan et al. 2006). The study has also expanded the evidence base for CAT within forensic populations (Pollock et al. 2006). The primary objective was to provide indications as to feasibility of delivering G-CAMT during routine care in a high security hospital, and not to provide a definitive test of the efficacy of G-CAMT on an a priori specified primary outcome measure (Thabane et al. 2010). We rather sought as a secondary objective to conduct an initial examination of the treatment effects of G-CAMT and compare those to outcomes achieved by TAU. Meeting these objectives would usefully signal whether a fully powered main RCT could and should then be conducted (Lancaster et al. 2010). This has been achieved with satisfying results regarding G-CAMT attendance, good results for satisfaction, the effect sizes produced by G-CAMT being encouraging and some initial differences with TAU being apparent in terms of clinical outcomes, on some analyses (albeit in small sample sizes and depending on time of measurement). No differences were apparent in between-arm comparisons in the standard care provided to each arm during the intervention, with participants having equal amounts of psychological input hours, physical activity hours, recreational music hours and hours spent in other activities. In future G-CAMT studies, the PROQ-2 (Birtchnell & Evans, 2004) would appear to be a suitable primary outcome measure, as the relational approach of the therapy and the measure appear

well matched.

Primary outcomes; feasibility, acceptance and satisfaction

In terms of ease of implementation of the intervention, G-CAMT was enabled through the delivery of a two-day training workshop in the CAT model and then supported via weekly clinical supervision. This indicates that for qualified music therapists, core CAT concepts were relatively easy to learn and integrate into their practice, when supported by the structure of a detailed treatment manual and ongoing clinical supervision. This mirrors other CAT evidence that manualised translations of the model into the work of professional groups can be achieved relatively easily, because of the high face validity and non-technical language of the CAT model (Meadows & Kellett, 2017). The manual was intended to support an intervention designed to be easily integrated into extant forensic MDT treatment pathways, as research should add value to, rather than disturb, core MDT work (NICE, 2009).

The satisfaction scores on the CSQ-8 would suggest that participants felt satisfied with G-CAMT as an intervention, with the mean scores being typical of healthcare satisfaction ratings (Attkisson & Greenfield, 2004). A single participant unfortunately dropped out of G-CAMT. The low dropout rate in this study however compares well with the low dropout rate for CAT that has been reported across adult mental health settings (Evans et al. 2016; Kellett et al. 2013). It is worth noting that Hannibal et al (2012) audit of dropout from MT illustrated that for the 27 participants (10 with schizophrenia and 17 with personality disorder) whom started treatment ($M = 18$ sessions), 3 dropped out (i.e. creating a dropout rate of 11%). In combination with the evidence presented of G-CAMT participants often attending the majority of sessions offered, this would suggest that G-CAMT appears an acceptable approach to use with male MDOs. The high sessional attendance results would echo the Talwar et al. (2006) findings concerning high sessional attendance rates at MT and

the extant CAT evidence concerning high sessional attendance rates in routine practice (Evans et al. 2016; Kellett et al. 2013). The duration and approach of G-CAMT being a structured and manualised 16-week treatment would be suggestive of a time-efficient and containing intervention. This would dovetail with the original ethos of CAT as being fit for purpose for application in highly pressured public services (Ryle et al. 2014).

Patient preferences

Patient preference trials are a rarity in MT. Gold et al. (2014) started a standard RCT in a prison population (*N*=113), before switching to a design with a patient preference element. The study had to be halted due to ongoing recruitment issues, which included rapid turnover in the prison population. The length of stay of the average MDO of 6-years in a high secure setting (Butwell et al. 2000) would not appear to present a recruitment challenge in terms of rapidity of turnover. There are however advantages to continuing to try to adopt this methodology in future forensic MT outcome research. Firstly, patient preference trials are able to recruit ambivalent or recalcitrant potential forensic participants who would not otherwise get recruited, due to their refusal to be randomized (Moffett et al. 1999). Secondly, the results from studies that incorporate patient preferences have greater external validity (Howard & Thornicoft, 2006) and therefore a potential greater influence on forensic MT practice. Results from the patient preferences in the current study suggest that MDOs do have strong preferences (i.e. N=16 expressed a treatment preference compared to only N=4 that chose to be randomised) and future MT outcome research should recognise this. It is acknowledged that there is a debate as to the validity of comparing randomised and non-randomised participants in such preference designs (MacLehose et al. 2005).

Secondary outcomes

Glorney et al. (2010) emphasised that MDO treatments needed to be cost-effective. In terms

of reducing length of stay in high secure hospitals the current study assessed residency at 2-year follow-up. There was no convincing evidence for G-CAMT differentially enabling participants to be assessed as fit to move to less secure (and hence less expensive) treatment centres. Interesting changes within and between the arms emerged from the individual and group level analyses on the PROQ-2. The full-scale PROQ-2 analysis showed a small effect size at follow-up for G-CAMT participants in terms of improved interpersonal functioning; in the TAU arm a minimal effect size occurred. This would suggest that G-CAMT enabled participants to recognise and then revise the reciprocal roles and associated procedural sequences (Ryle & Kerr, 2002) maintaining their interpersonal difficulties. G-CAMT therefore appeared to create a small effect of increasing the ability to begin to relate differently to others, in a population where established interpersonal deficits are at the heart of their offending (MoJ, 2012). Where change did occur in the TAU arm (e.g. such as the RCI baseline-termination analyses), then this may have been an effect of the ongoing treatment or possibly due to the research participation process itself (Godin, Sheeran, Conner & Germain, 2008).

When change was analysed on a case-by-case basis in G-CAMT a small proportion (N=2) participants made reliable baseline to termination reductions in suspiciousness and intimidation. However, at follow-up N=2 G-CAMT participants experienced a reliable deterioration in their suspiciousness. In the CAT model with PD in non-forensic PD populations, then structured follow-up over 6-months post-treatment is a key feature of the model (Ryle & Kerr, 2002). This raised the possibility of more structured follow-up sessions being integrated into the G-CAMT manual, to ensure scaffolding of changes made during the group having better durability over time. A recent further iteration of the G-CAMT manual has been the addition of structured follow-up (Compton Dickinson, 2017). CAT is characterised as a therapy that tries to recognise the emotional power of endings (Ryle &

Kerr, 2002), but the deteriorations in the ND (suspicious) and UD (intimidating) domains

may have been a reaction to G-CAMT termination. This demonstrates that passive music

listening appears to have little therapeutic effect in MDOs (Compton Dickinson, 2015).

Strengths of the current research

A strength of the current study has been its strong theoretical foundations that have then been

clearly translated into the group-based music therapy intervention, tailored to the needs and

treatment of MDOs. A previous review of N=22 MT outcome studies with oncology patients

(containing 18 RCTs) found that only two studies stated the theoretical framework that

guided intervention content (Burns, 2012). The future development of the G-CAMT model

could also learn from methodological and theoretical best practice approaches being

implemented within the cognitive-behavioural MT evidence base (Silverman, 2008).

Additionally, a recent review of the reporting quality of music intervention research in

healthcare (Burns et al. 2018) noted that of the N=187 studies meeting inclusion criteria, less

than 50% provided enough detail about the music intervention and so inhibiting interpretation

and cross-study comparisons. The detail of the description of G-CAMT provided in Table 1

would therefore be an example of a well-designed and described MT intervention.

Methodological critique and future methodological directions

In terms of study weaknesses, the most obvious issue concerns the small sample sizes in the

arms of the study. Guidelines suggest that the sample size of a pilot trial needs to be N=12 in

each arm (Thabane, Ma & Chu, 2010) and the current study fell short of this. From the pool

of potential participants (N=102), 31% refused to participate in the research citing reasons

related to being currently engaged in other recreational or therapeutic activities.   The aim

was to recruit a sufficient number of participants for a feasibility pilot trial and this (on

reflection) was naive to the impact that engaging in current recreational or therapeutic

activities might affect motivation to engage in the trial. Indeed, seven participants were willing to participate in the research and met inclusion criteria, but preferred to be allocated to treatment as usual (TAU). This possibly reflects participants currently engaged in high levels of MDT treatment. Four participants were willing to be randomised and therefore did not state any preference. It was not surprising that more participants stated a preference than not, and this is consistent with the issues presented by adopting a patient preference research method (Brocklehurst, 2005). It is important in the design of patient preference trials that the treatment choices are described in a balanced and considered manner. Patient information leaflets were reviewed and approved by the ethics committee, by clinicians and potential participants to ensure that the treatment options were presented in an equipoise manner in relation to content and approach.

The sample size limitation also applies in relation to the study being a designed as a patient preference trial (Awad et al. 2000), as this choice of outcome methodology is reliant on large sample sizes (i.e. similar to non-inferiority trials) and this was not the case here. Sufficient recruitment to power trials in high secure hospitals that can reliably detect differences between arms is challenging due to the relatively smaller size of the patient population suitable for treatment and also the intensity of, and commitment to on-going MDT treatment (Shepherd, Boardman & Slade 2008). It may have been scientifically premature (on reflection) to attempt a patient preference trial, but it is worth noting that ethical approval was granted on condition that this population of incarcerated patients (who have very limited choices per se) could feel sufficiently empowered to choose their treatment (or not and be happy to be randomised). A better scientific option may have been a feasibility two-arm parallel group design, or a three-arm parallel group design in which one of the arms was a generic MT intervention. The plans for such a study are in the early developmental stages. Setting an alpha of .05 and power of 0.80 and given the meta-analysed effect size of 0.84 for

CAT (Ryle et al. 2014), then a sample size of N=48 would be required (i.e. 24 participants in each arm) using the PROQ-2 as the primary outcome measure. It may however be reasonable to expect a medium between group effect size. This would in all likelihood demand a multi-centre approach, which would be feasible across the secure hospitals and penitentiaries, given the acknowledged issues with recruitment of MDOs (Shepherd, Boardman & Slade 2008). If such a study were to show that G-CAMT differentially improved relational abilities, then the relational components of the model (in comparison to generic MT and TAU) would be the assumed mechanism of change. Future studies also need to include economic evaluations of clinical interventions, as Glorney et al. (2009) highlighted the need for cost-effectiveness to be considered alongside clinical effectiveness (e.g. particularly in terms of reduced stay in high secure hospitals). Therefore, clinical and organisational relevant outcomes would also be reductions to risk and an associated move to lower security (and low cost) setting. The medication participants were receiving was not recorded and the sample was restricted (due to context) to male MDOs. The effectiveness of G-CAMT with females was reported in an early stage G-CAMT pilot project (Lawday & Compton, 2013).

The follow-up period for the primary outcome measure was short and would have benefited from a genuine long-term follow-up period (i.e. matching the residency assessment follow-up at 2-years). Whilst it was an advantage that the assessor who collected the self-report psychometric outcomes was masked to allocation, it was impossible to mask the nursing staff who completed the CIRCLE measure. In terms of measurement of dissociation used (i.e. the MDI; Briere, 2002), it may have been more theoretically consistent with a CAT intervention to use the Personality Structure Questionnaire (Pollock et al, 2001). The PSQ would be a useful secondary measure in any future trial, particularly given the PSQs ability to detect personality disorder (Berrios, Kellett, Fiorani & Poggioli, 2016). Whilst adverse effects were considered in the ethical approval for the study, future studies should consider

and record any adverse effects in detail. Whilst there is a measure of competency of CAT delivery for the one-to-one approach (Bennett & Parry, 2004), there is currently no competency measure developed for CAT delivered in groups or via MT. Whilst adherence was aided with the development and use of a treatment manual, use of a treatment adherence measure would have improved the internal validity of this study. The issue of treatment integrity (i.e. considering both adherence to a manualised intervention and competence in terms of delivery; Perepletchikova, Treat & Kazdin, 2007) should therefore be considered in detail in future G-CAMT research. Therefore, any future G-CAMT trial needs to consider using a three-arm parallel group design, changing the secondary measures, assessing adherence/competency, recording adverse effects and extending the follow-up period.

Conclusions

In summary, this has been the first feasibility quasi-experimental patient preference study of MT with an MDO population in a secure hospital setting. The results mirror the evidence base for MT, which as an adjunct treatment for schizophrenia compared to TAU (i.e. a high secure MDT treatment routine in the current context), then the addition of MT into care plans of MDOs appears useful (Tseng, et al. 2016). This work chimes with the call to provide theory-based interventions within psychiatric MT that are then tested within well-controlled outcome methodologies (Silverman, 2015). This study has particularly developed the evidence base for MT by evaluating adding in an evaluation of relationality and also evaluating an avowedly relational MT model. G-CAMT appears to offer a grounding theory, therapeutic relationship, tools and techniques that make it possible to consider and change relational dynamics rooted in and mirroring early developmental traumas (Ryle & Kerr, 2002), within a structured and containing short-term group environment. Future fully powered and theoretically informed main clinical trials using both passive (i.e. TAU) and/or active (e.g. cognitive behaviourally informed MT; Hakvoort, 2014) controls with

comparisons of outcomes across the genders in high secure hospitals now appears indicated.

Table 1; description of the G-CAMT phases and associated CAT tools						
Stage one: sessions 1-4	**Purpose of sessions**	**Content and structure**	**Suggested musical resources**	**CAT tools**	**Therapeutic aims and techniques**	**Homework tasks**
Recognition Tasks **Establish safety and basic G-CAMT concepts**.	• Establish a positive therapeutic rapport. • Define group boundaries. • Orientate the group to the present.) • Assess whether the patient can choose to actively engage with the instruments. • Ensure the patient is not overwhelmed or intimidated. • Explain the purpose of therapy • Explain collaboration	• Introduce turn-taking in dialogue. • Introduce the 'safety call' • Reflect on: 'What is;' *rather than* 'if only' or 'what was.' • Start with breathing exercises and talk through a short body scan grounding visualisation. • Observation and description of instruments. • Introduce a limited range of instruments to facilitate mindfulness and group harmony. • Encourage mild curiosity to extend sensory experience through sound, touch and visual qualities.	• Tibetan Singing bowl and/or Tibetan chimes. • Djembe drum. sounding bowl • Hand held percussion. • Power chimes, • Bass Bar for grounding. • No large instruments that break the group circle. • Introduce patients to the 'Sounding Bowl' (a strung instrument, but unlike a guitar, so no pre-conceptions or fantasies of fame are elicited which otherwise bring out narcissistic traits.) • Build awareness of the different senses, from the visual/tactile/ vibrational/auditory qualities	• Psychotherapy File. • States descriptions and self-states.' • Assess most appropriate relational patterns • Explain and draw the concept of 'reciprocal role relating.'	• Attend to one sensory experience at a time and introduce the 'Sound Print' • Build up embodied recognition of touch/visual impression/sound/ smell & auditory responses; leading to safe exploration when starting musical engagement. • Reiterate no judgement. • Promote active positive engagement, unconditional positive regard. • Encourage choice without pressure to touch or play the instruments. • Assess whether the patient may feel overwhelmed with sensory input or fearful of exposure/humiliation.	• Ask the patient to think about what they would like to achieve • Define target problem. • Self- reflection on acceptance of what has occurred in life-path,

		• Encourage active verbal engagement in prompted dialogue towards sharing to shared reciprocity.				
Stage two: sessions 5-8	**Purpose of the sessions**	**Content and structure**	**Suggested musical resources**	**CAT tools**	**Therapeutic aims and techniques**	**Homework tasks**
Engagement with feelings of self and others	• Encourage patients to engage with the feelings within the group's music and on a personal level. • Development of connection to and empathy for others	• Short, emotionally attuned improvisations that come out of the verbal dialogue. • Prompt musically rather than allow extended musical narrative. • Use relevant daily issues and themes in the here and now. • Ensure music is harmonious and aesthetically pleasing rather than chaotic. • As choice develops; introduce the concept of the	• Music Therapist: increased use of singing and vocalizing. • Use the 'Sounding Bowl' • Bass Bar for grounding and mediating affect. • Temple blocks and Congas for wider dynamic range and to play whilst moving around and standing up. • Encourage exploration with supportive prompting on choices	• Reinforce the concept of, and name the emergence of reciprocal roles in musical and verbal dialogue. • Start to formulate a skeletal mapped diagram of group relating in supervision. • Introduce a flip chart and coloured pens in the session for group relational map drawing to	• Find commonalities. • Recognition of each patient's reciprocal relating procedures (RRPs) as enacted within the group (e.g. controlling-controlled, dominating-dominated). • This leads to recognition of the target problem procedure (developed from the target problem. • Once observed and named this constitutes verbal reformulation.	• Journal keeping to develop self-reflection • Promote recognition of when the individual's target problem emerges in daily interactions

				enhance cognitive recognition and understanding.		
		'Sound print', to promote individual identifications with, and sharing of instruments.				
Stage Three: sessions 9-12	**Purpose of the session**	**Content and Structure**	**Suggested musical resources**	**CAT tools**	**Therapeutic aims and techniques**	**Homework tasks**
Tolerating interpersonal differences and exploring conflict resolution **(Revision)**	• Managing impulsivity. • Cognitive and emotional awareness of rage responses • Recognition of 'the dark side' in human nature and the need for altruism. • Therapist contains, role models and mediates negotiations towards understanding and accepting	• Widen the range of emotional expression so that anger and frustration can be safely expressed and acknowledged through a wider range of dynamics and harmonies and dissonances. • Explain the difference between, and impact of assertiveness (demanding their needs to be met) and aggression as	• Increased range of chromatically tuned instruments and potential for dissonant harmonies • Include larger Instruments: Piano, Guitar, Tam Tam Temple blocks, keyboard. • Music Therapist uses their specialist instrument to contain strong emotions.	• Psychotherapy File: CAT Traps and Dilemmas • State Description Procedure to reinforce presence of both victim and offender self-states in all individuals-. • Ideally set up a table and chairs separate to the musical circle to collaboratively sit and draw/	• Develop energetic, cathartic input for expression of anger and release of negativity. (Including, fear of the humiliating to humiliated RR.) • MT to consider appropriate timing for harmonic resolutions to dissonant and abrasive sounds. • Hold on to aesthetic musical qualities • Manage the potential of over-arousal or offence paralleling behaviors • Use repeating cadences as /if required to facilitate harmonious endings.	• Individuals to recall an ordinary issue of disagreement and record how they dealt with it in the past and how to they may do something differently in the present. • Recognize when passive aggressive • Walking away as a safety procedure to calm down. • Recognize dilemma: Either I feel all bottled up Or making a terrible mess • Question: Owning feelings of

	differences of opinion and conflicts enacted musically within the group. • Encouraging verbal debate and dialogue and self-expression • Facilitate verbal recognition of negativity following improvisation.	a violent behaviour.	• Call and response drumming in dialogue.	map these identified self-states together.	• Freer rhythmic work, dynamic range. (Aim: to explore the impact of very loud sounds on others and self.)	resentment? Sharing- shared talk with named nurse.
Stage Four: sessions 13-16	**Purpose of sessions**	**Content and structure**	**Suggested musical resources**	**CAT tools**	**Therapeutic aims and techniques**	**Homework tasks**
Preparing for closure. **Saying farewell** **Risk Assess patient well-being prior to all**	• Management of the emergence of separation anxieties and the re-emergence of offence – related behaviors,	• Extend emotionally related improvisations with improved collaboration in group negotiation of how to begin and how to	• Music Therapist to increase or modify resources weekly according to session content. • Use the full range of instruments if risk assessed as safe.	• Recognise and acknowledge the re-emergence of reciprocal roles such as abandoning-abandoned, rejecting-rejected,	• Reflection and preparation of how each individual chooses to say 'good-bye' and to safely separate from others. • Music Therapist to write and share within the group, farewell	• Farewell letter writing or musical gesture planning. • Journal keeping of own changes in ability to help others to accept help.

sessions, review clinical notes for any incidents during the previous week.	and the psychological defenses of denial of sad or angry feelings or acceptance • Internalize other to self-roles 'having been heard- to ability to hear others. • Acceptance: 'it is okay to feel sad' feelings do not kill us and they will pass. • Explore plans and potential for moving on to new beginnings. • Name and address distress, aim to allow the group to ameliorate and help each other.	resolve the ending of each piece. • Focus on sharing to shared responsibility. (Symbolically each individual is therefore negotiating non-verbally, and connecting to others in the music.) • Begin each musical improvisation from silence and end collaboratively in harmony and explore different forms of silence, • Observe for feelings of remorse; desire to make amends for wrongdoing.	• Hold on to aesthetic musical qualities • Manage the potential of over-arousal or offence paralleling behaviors • Use repeating cadences as /if required to facilitate harmonious endings. • Provide solace with shared playing (e.g. of sounding bowl) MT vocalize to accompany. .	annihilating-annihilated. • Complete the group SDR together. • Consolidate *"scaffolded learning"* (Vygotsky) • Agree individual packages of therapy tools: Psychotherapy file, Individual Target Problem. Copies of the Group SDR and fare-well letters.	letters to each member naming their individual contributions. • Offer the group members the choice to write their own farewell letters to share with the whole group or to plan appropriate musical gestures to represent the same. • Final session: individual letters and/or fare-well gestures are shared in the group: • Celebrate achievement in commitment to completing the course of G-CAMT treatment.	

Table 2. Sample demographic and clinical/forensic characteristics

Characteristic	Total (n = 16)	G-CAMT (n = 10)	TAU (n = 6)	U value	*P*
Age in years (mean, SD)	38.81 (11.21)	38.70 (9.38)	39.00 (14.79)	29.50	0.96
Ethnicity (%)					
White British	12	7	5		
Black/Black British	2	1	1		
Asian/Asian British	1	1	0		
Other ethnic group	1	1	0		
Mental health diagnosis (%)					
Schizophrenia	9	8	3		
Psychopathic	1	1	0		
Other, psychotic	4	1	3		
Co-morbid personality Disorder	2	1	1		
Additional factors (%)					
Attempted suicide: Yes	9	7	2		
Substance misuse	13	10	3		
Index offence (%)					
Murder	5	3	2		
Manslaughter	3	3	0		
Bodily harm	4	1	3		
Sexual assault	2	1	1		
Arson (%)	1	1	0		
Attempted murder	1	1	0		
Forensic details (mean, SD)					
Age at first conviction, yrs	22.19 (9.35)	21.70 (7.85)	23.00 (12.25)	26.50	0.96
Length of offending in yrs	15.00 (10.07)	14.70 (9.32)	15.50 (12.14)	28.50	0.88
Length of incarceration, yrs	5.80 (4.70)	6.08 (5.33)	5.33 (3.84)	19.50	0.38
Measures (mean, SD)					
PROQ2	144.56 (28.98)	148.10 (25.69)	138.67 (35.57)	24.00	0.56
MDI	45.13 (14.00)	46.10 (16.79)	43.50 (8.69)	29.50	0.96
BES	67.63 (8.02)	66.00 (8.04)	70.33 (7.89)	22.00	0.43
CIRCLE: hostile	5.25 (2.93)	6.33 (2.55)	2.00 (0.00)	2.00	0.006*
CIRCLE: withdrawn	7.83 (3.81)	7.56 (4.36)	8.67 (1.53)	11.50	0.73
CIRCLE: sociable	7.58 (2.47)	7.67 (2.83)	7.33 (1.15)	4.00	1.00
CIRCLE: friendly	11.33 (3.73)	10.22 (3.53)	14.67 (2.08)	13.50	1.00

Note. *p<.05

Table 3: Individual participant reliable improvement and deterioration rates based on intention-to-treat data on the PROQ-2 sub and full-scale scores[a]

	G-CAMT	TAU	G-CAMT	TAU
	n	n	n	n
Improvement[b]				
Full-scale score	0	2	1	2
Upper Neutral	0	2	0	0
Upper Close	0	0	2	1
Neutral Close	0	0	2	0
Lower Close	0	1	0	0
Lower Neutral	0	1	0	1
Lower Distant	0	0	0	0
Neutral Distant	0	2	0	2
Upper Distant	0	2	0	1
Deterioration[b]				
Full-scale score	0	1	0	0
Upper Neutral	0	0	0	0
Upper Close	0	0	0	0
Neutral Close	0	0	0	0
Lower Close	0	0	0	1
Lower Neutral	0	1	0	1
Lower Distant	0	1	0	1
Neutral Distant	2	0	2	0
Upper Distant	0	0	1	0

[a] Findings are based on the number of participants (including dropouts) for whom reliable change was statistically possible at each time-point. [b] Reliable change was indicated by the following change scores: Full Scale = 29.62, UN = 7.1, UC = 8.0, NC = 7.9, LC = 6.9, LN = 7.1, LD = 8.1, ND = 6.8, and UD = 7.3.

Table 4: Intention to treat PROQ-2 scores over time, within-arm effect size estimates and between-arm comparisons at post-treatment and follow-up

| | Within-arm Analyses | | | | | | Between-arm Analyses |
| | G-CAMT (N = 10) | | | TAU (N = 6) | | | G-CAMT vs. TAU change scores |
Measure and time point	*M*	*SD*	Effect size[a] (95% CI)	*M*	*SD*	Effect size[a] (95% CI)	*U (p)*
Upper Neutral							
Pre	18.20	4.73		19.83	4.96		
Post	18.80	5.33	-0.13 (-0.10-0.76)	16.50	6.98	0.67 (-0.54-1.78)	27.50 (.79)
FU	18.30	5.81	-0.02 (-0.90-0.86)	17.00	4.69	0.57 (-0.62-1.68)	21.00 (.37)
Upper Close							
Pre	15.50	6.62		12.17	5.38		
Post	13.60	6.95	0.29 (-0.61-1.16)	12.50	5.47	-0.06 (-1.11-1.08)	17.50 (.18)
FU	11.10	6.69	0.66 (-0.26-1.53)	9.67	4.97	0.46 (-0.72-1.57)	19.00 (.26)
Neutral Close							
Pre	21.30	7.65		20.17	6.70		
Post	20.80	7.19	0.07 (-0.81-0.94)	20.00	5.10	0.03 (-1.11-1.16)	30.00 (.30)
FU	18.20	8.46	0.41 (-0.50-1.27)	19.83	6.18	0.05 (-1.08-1.18)	19.50 (.26)
Lower Close							
Pre	21.40	5.56		21.00	4.98		
Post	20.40	6.08	0.18 (-0.71-1.05)	20.17	7.36	0.17 (-0.98-1.29)	24.00 (.56)
FU	21.60	5.93	-0.04 (-0.91–0.84)	21.50	3.27	-0.10 (-1.22-1.04)	29.50 (.96)
Lower Neutral							
Pre	19.90	6.95		16.17	8.08		
Post	16.60	6.59	0.47 (-0.43-1.34)	16.38	7.57	-0.03 (-1.16-1.11)	14.50 (.09)
FU	16.60	6.59	0.47 (-0.43-1.34)	16.67	7.15	-0.06 (-1.19-1.07)	14.50 (.09)
Lower Distant							
Pre	19.50	7.92		16.33	8.87		
Post	18.30	8.03	0.13 (-0.76-1.00)	18.83	9.16	-0.29 (-1.40-0.87)	18.50 (.22)
FU	17.40	7.68	0.27 (-0.63-1.13)	19.00	6.99	-0.30 (-1.42-0.86)	30.00 (1.0)
Neutral Distant							
Pre	13.30	6.29		16.00	7.16		
Post	16.70	5.17	-0.54 (-1.41-0.37)	12.17	6.56	0.53 (-0.66-1.64)	7.00* (.01)
FU	16.30	6.25	-0.48 (-1.34-0.43)	11.17	7.05	0.67 (--0.54-1.78)	6.00* (.007)
Upper Distant							
Pre	19.00	5.16		18.16	2.32		
Post	20.00	5.35	-0.19 (-1.06-0.69)	15.83	4.26	1.00 (-0.26-2.12)	20.50 (.30)
FU	18.80	5.61	0.04 (-0.84-0.91)	17.50	3.73	0.28 (-0.87-1.40)	30.00 (1.0)
Full scale							
Pre	148.10	25.69		138.67	35.57		
Post	145.20	26.51	0.11 (-0.77-0.99)	132.83	21.87	0.16 (-0.78-1.28)	27.00 (.79)
FU	138.30	20.70	0.38 (-0.52-1.25)	132.50	23.65	0.17 (-0.97-1.29)	20.00 (.31)

Note. *p <0.025 (.05/2 Bonferroni correction due to multiple comparisons)[a] and within-group effect size estimates calculated as pre-treatment mean minus the post-treatment mean, divided by the pre-treatment SD

Table 5: Intention to treat secondary outcome scores over time, within-arm effect size estimates and between-arm comparisons at post-treatment and follow-up

	Within-arm Analyses						Between-arm analyses
	G-CAMT (N = 10)			TAU (N = 6)			G-CAMT vs. TAU change scores
Measure and time point	M	SD	Effect size[a] (95% CI)	M	SD	Effect size[a] (95% CI)	U (p)
BES[b]							
Pre	66.00	8.04		70.33	7.89		
Post	69.00	12.70	-0.37 (-1.24-0.53)	73.33	8.64	-0.38 (-1.49-0.79)	27.50 (.79)
FU	69.50	9.79	-0.44 (-1.3-0.47)	75.83	7.31	-0.70 (-1.80-0.52)	22.00 (.43)
MDI							
Pre	46.10	16.79		43.59	8.69		
Post	49.20	18.67	-0.18 (-1.06-0.70)	48.17	7.19	-0.53 (-1.63-0.66)	24.00 (.56)
FU	46.90	17.24	-0.05 (-0.92-0.83)	44.50	9.69	-0.10 (-1.23-1.04)	29.00 (.96)
CIRCLE Sociable[b]							
Pre	7.67	2.83		7.33	1.15		
Post	8.78	3.46	-0.39 (-1.26-0.51)	6.67	2.31	0.57 (-0.62-1.68)	8.50 (.37)
FU	9.11	3.86	-0.51 (-1.38-0.40)	7.33	1.53	0.00 (-1.13-1.13)	12.00 (.86)
CIRCLE Hostile							
Pre	6.33	2.55		2.00	0.00		
Post	5.44	3.24	0.35 (-0.55-1.22)	3.60	2.70	-	7.50 (.28)
FU	4.56	2.96	0.69 (-0.24-1.57)	5.00	2.65	-	2.00 (.36)
CIRCLE Withdrawn							
Pre	7.56	4.36		8.67	1.53		
Post	6.89	3.75	0.15 (-0.73-1.02)	7.00	3.54	1.09 (-0.19-2.21)	8.00 (.37)
FU	7.67	3.74	-0.03 (-0.90-0.85)	8.33	2.08	0.22 (-0.93-1.34)	13.00 (1.00)
CIRCLE Friendly[b]							
Pre	10.22	3.53		14.67	2.08		
Post	11.00	3.00	-0.22 (-1.09-0.67)	9.20	4.32	2.63 (0.94-3.92)	3.50 (.06)
FU	11.89	3.30	-0.47 (-1.34-0.43)	10.67	2.89	1.92 (0.44-3.11)	0.50* (0.009)

Note. *p <0.025 (.05/2 bonferoni correction due to multiple comparison)

BES = Basic Empathy Scale; MDI = Multi-Scale Dissociation Inventory; CIRCLE = The Chart of Interpersonal Reactions in Closed Living Environments.

[a] Within-group effect size estimates calculated as pre-treatment mean minus the post-treatment mean, divided by the pre-treatment standard deviation.

[b] Effect sizes with a negative sign denote positive treatment effects for this measure.

Figure 1. Consort flow chart of preference, treatment and analysis
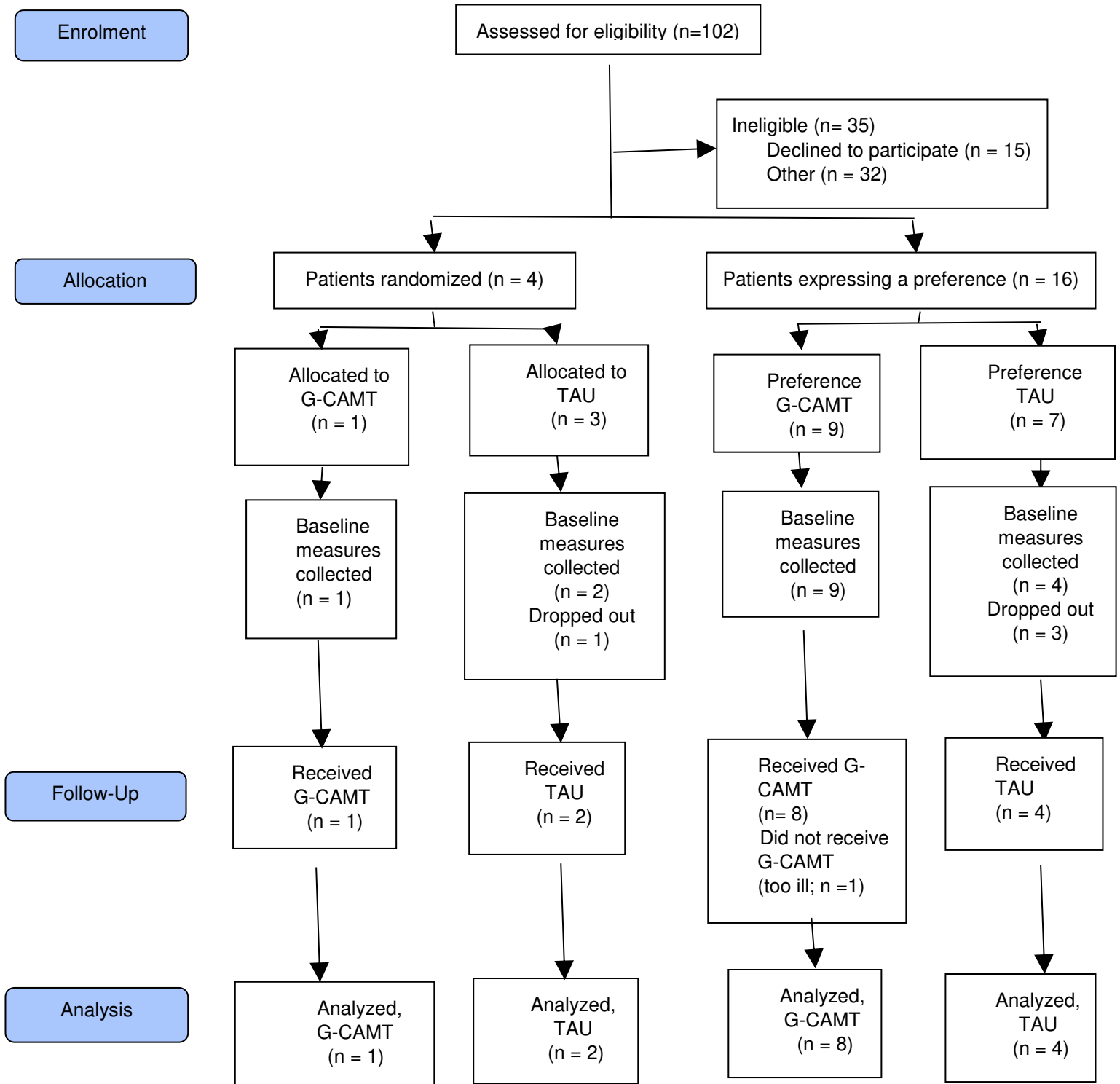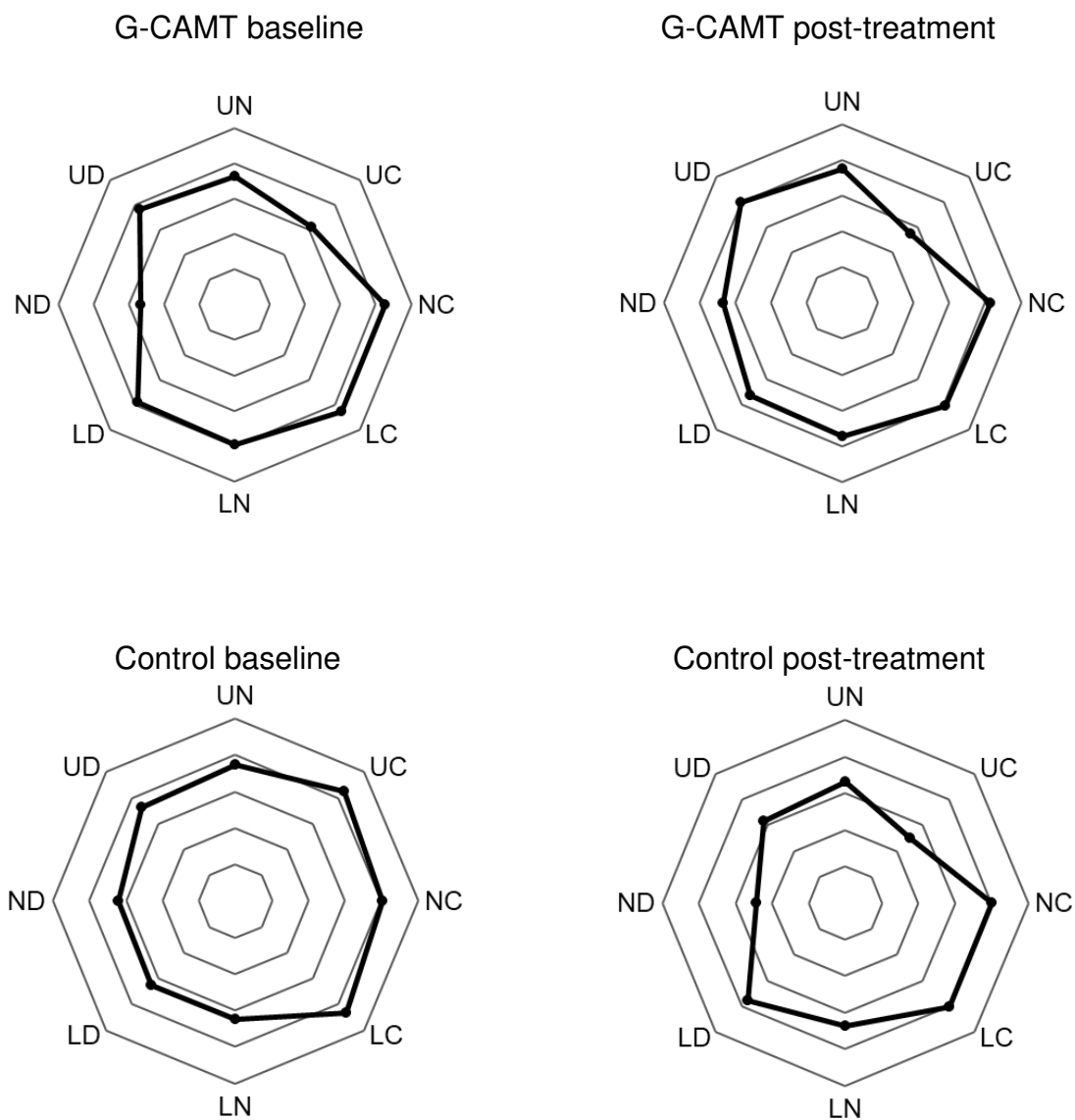
Figure 2. PROQ-2 baseline to termination domain scores* in G-CAMT and TAU

G-CAMT baseline

G-CAMT post-treatment

Control baseline

Control post-treatment

*scores closer to the centre of the octagon indicate lower scores on PROQ-2 domains and so improved outcomes and vice-versa

References

Attkisson, C. & Greenfield, T. (1994). Satisfaction Questionnaire-8 and the Service

Satisfaction Scale-30. In *Psychological Testing: Treatment Planning and Outcome*

Assessment. San Francisco, CA: Lawrence Erlbaum Associates.

Attkisson, C.C., & Greenfield, T.K. (2004). The UCSF Client Satisfaction Scales: I. The

Client Satisfaction Questionnaire-8. In M. Maruish (Ed.) *The use of psychological testing for*

treatment planning and outcome assessment (3rd Ed.). Mahwah, NJ: Lawrence Erlbaum

Associates.

Awad, M.A., Shapiro, S.H., Lund, J.P., & Feine, J.S. (2000). Determinants of patients'

treatment preferences in a clinical trial. *Community Dentistry and Oral Epidemiology, 28,*

119-125. doi: 0.1034/j.1600-0528.2000.028002119.x

Bakhtin, M.M. (1981). *The dialogic imagination: four essays*. Michael Holquist (Ed). Austin

and London: University of Texas Press.

Bennett, D., & Parry, G. (2004). A measure of psychotherapeutic competence derived from

cognitive analytic therapy. *Psychotherapy Research,* 14, 176-192.

doi.org/10.1093/ptr/kph016

Berrios, R., Kellett, S., & Fiorani, C. & Poggioli, M. (2016). Assessment of identity

disturbance: factor structure and validation of the Personality Structure Questionnaire in an

Italian sample. *Psychological Assessment*, 28 (4). doi: 10.1037/pas0000143.

Birtchnell, J., Falkowski, J., & Steffert, B. (1992). The negative relating of depressed

patients: A new approach. *Journal of Affective Disorders, 24*, 165-176.

doi.org/10.1016/0165-0327(92)90064-D

Birtchnell, J., Denman, C., & Okhai, F. (2009). Cognitive analytic therapy: comparing two measures of improvement. *Psychology and Psychotherapy: Theory, Research and Practice*, *77*, 479–92. doi: 10.1348/1476083042555398

Birtchnell, J., & Evans, C. (2004). The persons relating to others questionnaire. *Personality and Individual Differences*, *36*, 125-40. doi: org/10.1016/S0191-8869(03)00073-4

Birtchnell, J., & Shine, J. (2000). Offence typology and the interpersonal octagon: an exploratory analysis. *Criminal Behaviour and Mental Health*, *2*, 227-29. doi: 10.1002/cbm.498

Birtchnell, J., & Shine, J. (2000). Personality disorders and the interpersonal octagon. *British Journal of Medical Psychology, 73,* 433-448.

Birtchnell, J., Shuker, R., Newberry, M. & Duggan, C. (2009). An assessment of change in negative relating in two male forensic therapy samples using the persons relating to others questionnaire. *Journal of Forensic Psychiatry and Psychology, 20*, 387–407. doi: org/10.1080/14789940802542840

Blackburn R. (1992). Criminal behaviour, personality disorder and mental illness: the origins of confusion. *Criminal Behaviour and Mental Health, 2*, 66-77. doi:10.1002/cbm.1992.2.2.66

Blackburn, R., & Renwick, S. (1996). Rating scales for measuring the interpersonal circle in forensic psychiatric patients. *Psychological Assessment*, *8*, 76–84. doi: 10.1037/1040-3590.8.1.76.

Blackburn R., & Glasgow D. V. (2006). Manual for the Chart of Interpersonal Reactions in Closed Living Environments (CIRCLE). Unpublished manuscript, University of Liverpool.

Bowen, D.J., Kreuter, M., Spring, B., Cofta-Woerpel, L., Linnan, l., Weiner, D., Bakken, S., Kaplan, C.P., Squiers, L., Fabrizio, C., & Fernandez, M. (2010). How we design feasibility

studies. *American Journal of Preventative Medicine, 36,* 452–457.

doi: 10.1016/j.amepre.2009.02.002

Briere, J. (2002). *Multiscale dissociation inventory professional manual*. Odessa, FL:

Psychological Assessment Resources.

Briere, J., Weathers, F.W., & Runtz, M. (2005). Is dissociation a multidimensional construct?

Data from the Multiscale Dissociation Inventory. *Journal of Traumatic Stress*, *18*, 221 - 231.

doi: 10.1002/jts.20024

Brocklehurst, P. (2005). Partially randomised patient preference trials. *British Journal of*

Obstetrics and Gynaecology, 104, 1332-1335. doi.org/10.1111/j.1471-0528.1997.tb10999.x

Butwell, M., Jamieson, E., Leese, M., & Taylor, P.J. (2000). Trends in special (high-security)

hospitals. 2: Residency. *British Journal of Psychiatry, 176,* 260-265. doi:

10.1192/bjp.176.3.260

Burns, D. (2012). Theoretical rationale for music selection in oncology intervention research:

an integrative review. *Journal of Music Therapy, 49*, 7-22.

Carpenter, J.R., & Kenward, M.G. (2008). Missing data in clinical trials - a practical guide.

Birmingham: National Institute for Health Research; 2008. Publication RM03/JH17/MK.

Available   at

http://www.pcpoh.bham.ac.uk/publichealth/methodology/projects/RM03_JH17_MK.shtml.

Carr, C., Odell-Miller, H., & Priebe, S. (2013). A systematic review of music therapy practice

and outcomes with acute adult psychiatric in-patients. *PLoS One, 2*;8(8). doi:

10.1371/journal.pone.0070252

Calvert, R., & Kellett, S. (2014). Cognitive analytic therapy: A review of the outcome evidence base for treatment. *Psychology and Psychotherapy: Theory, Research and Practice, 87*, 253-77. doi: 10.1111/papt.12020

Campbell, M., Fitzpatrick, R., & Haines, A. (2000). Framework for design and evaluation of complex interventions to improve health. *British Medical Journal*, *321*, 694–696. doi: https://doi.org/10.1136/bmj.321.7262.694

Campbell, N.C., Murray, E., Darbyshire, J., Emery, J. *et al.* (2007). Designing and evaluating complex interventions to improve health care. *British Medical Journal*, *334*, 455– 459. doi: https://doi.org/10.1136/bmj.39108.379965.BE

Chen, X.J., Leith, H., Aarø, L.E., Manger, T., & Gold, C. (2016). Music therapy for improving mental health problems of offenders in correctional settings: systematic review and meta-analysis. *Journal of Experimental Criminology*, 1–20. doi: 10.1007/s11292-015-9250-y

Cohen, J. (1992). A power primer. *Psychological Bulletin, 112**, 155–159.*

Compton Dickinson, S. J. (2015). *A feasibility trial of group cognitive analytic music therapy in secure hospital settings.* PhD. Anglia Ruskin University.

Compton Dickinson, S.J., & Hakvoort, L. (2017). *The clinician's guide to forensic music therapy: Treatment manuals for group cognitive analytic music therapy and music therapy anger management.* London, UK: Jessica Kingsley Publishers.

Connell, J., Barkham, M., & Mellor-Clark, J. (2008). The effectiveness of UK student counselling services: an analysis using the CORE system. *British Journal of Guidance & Counselling, 36,* 1-18. doi: org/10.1080/03069880701715655

Duggan, C., Adams, C., McCarthy, L., Fenton, M. *et al.* (2006) *A systematic review of the effectiveness of pharmacological and psychological treatments for those with personality disorder*. Retrieved from www.nfmhp.org.uk/MRD%2012%2033%20Final%Report.pdf [Accessed 21 July 2014]

Evans, C., Margison, F., & Barkham, M. (1998). The contribution of reliable and clinically significant change methods to evidence-based mental health. *Evidence-Based Mental Health*, 1, 70-72. doi: 10.1136/ebmh.1.3.70

Evans, C., Mellor-Clark, J., Margison, F., Barkham, M., Audin, K., Connell, J., & McGrath, G.C. (2000). CORE: clinical outcomes in routine evaluation. *Journal of Mental Health*, *9*, 247-255. doi: org/10.1080/jmh.9.3.247.255

Evans, M., Kellett, S., Heywood, S., Hall, J., & Majid, S. (2017). Cognitive analytic therapy for bipolar disorder: A pilot randomized controlled trial. *Clinical Psychology and Psychotherapy,* 24, 22-35. doi: 10.1002/cpp.2065

Fassaert, T., Segeren, M., Grimbergen, C., Tuinebergen, W., & de Wit, M. (2016). Violent offenders as a target population for Public Mental Health Care. *Journal of Forensic and Legal Medicine, 40*, 54-57. doi: 10.1016/j.jflm.2016.03.003.

Gilligan, (2015). Editorial. *International Journal of Applied Psychoanalytic Studies 12,* 89-95.

Godin, G., Sheeran, P., Conner, M., & Germain, M. (2008). Asking questions changes behavior: Mere measurement effects on frequency of blood donation. *Health Psychology, 27*, 179-184. doi: 10.1037/0278-6133.27.2.179.

Gold, C., Assmus, J., Hjørnevik, K·, Qvale, L.G., Brown, F,K., Hansen, A.L., Waage, L· &

Stige, B. (2014). Music therapy for prisoners: pilot randomised controlled trial and

implications for evaluating psychosocial interventions. *International Journal of Offender*

Therapy and Comparative Criminology, 58, 1520-39. doi: 10.1177/0306624X13498693

Glorney, E., Perkins, P., Adshead, G., McGauley, G.A., Noak, J., Murray, K., & Sichau, G.,

(2009). Reducing length of stay in high secure hospital settings: a model for streamlining care

in Broadmoor Hospital. *The International Journal of Forensic Mental Health*, *9*, 138-48. doi:

org/10.1080/14999013.2010.499552

Glyn, J. (2009). Two's company, three's a crowd. In: H. Odell-Miller and E. Richards (Eds.)

2009. *Supervision of music therapy: a theoretical and practical handbook*. London, UK:

Routledge.

Green, K., Brown, K., & Chew, S. (2017). The relationship between childhood maltreatment

and violence to others in individuals with psychosis: a systematic review and meta-analysis.

Trauma, Violence and Abuse, 1-16. doi: 10.1177/1524838017708786

Jacobson, N.S., & Truax, P. (1991). Clinical significance: A statistical approach to defining

meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology,*

59, 12-19.

Jolliffe, D., & Farrington, D.P. (2006a). Development and validation of the Basic Empathy

Scale. *Journal of Adolescence*, *29*, 589–611. doi: 10.1016/j.adolescence.2005.08.010

Joliffe, D. & Farringdon, D.P. (2006b). The basic empathy scale in adults (BES-A); factor

structure of a revised form. *Psychological Assessment, 25,* 679-691. dio: 10.1037/a0032297

Hakvoort, L. (2014). Multifaceted music therapy in forensic psychiatry. In: J. de Backer and

Sutton. J [Eds.] *The music in music therapy- psychodynamic music therapy in Europe:*

clinical, theoretical and research approaches. London, UK: Jessica Kingsley Publishers.

Hannibal, N., Nygaard Pedersen, I., Hestbæk, T., Egelund Sørensen, T., Munk-Jørgensen, P. (2012). Schizophrenia and personality disorder patients' adherence to music therapy. *Nordic Journal of Psychiatry, 64,* 1-4. doi: 10.3109/08039488.2012.655775

Howard, L., & Thornicroft, G. (2006). Patient preference randomised controlled trials in mental health research. *British Journal of Psychiatry*, *188*, 303-304.

Kamioka, H., Kiichiro, T., Yamada, M., Park, H., Okuizumi,H., Tsuruoka,K., Honda, T., Okada, S., Park, S-J., Kitayuguchi, J., Abe,T. Handa, S., Oshio,T., & Mutoh, Y. (2014). Effectiveness of music therapy: a summary of systematic reviews based on randomized controlled trials of music interventions. *Patient Preference and Adherence, 8*, 727–754. doi: 10.2147/PPA.S61340

Kellett, S. (2005). The treatment of Dissociative Identity Disorder with Cognitive Analytic Therapy: Experimental evidence of sudden gains. *Trauma and Dissociation, 6, 55-81*. doi:10.1300/J229v06n03_03

Kellett, S. Bennett, D. Ryle, A. & Thake, A. (2013). Cognitive Analytic Therapy for Borderline Personality Disorder: Therapist competence and therapeutic effectiveness in routine practice. *Clinical Psychology and Psychotherapy, 20,* 216-225. doi: 10.1002/cpp.796

Kellett, S. Wilbram, M., Davis, C. & Hardy, G. (2014). Team consultancy using cognitive analytic therapy: a controlled study in assertive outreach. *Journal of Psychiatric and Mental Health Nursing, 21*, 687-97. doi: org/10.1111/jpm.12123

Kelly, P.J., Kyngdon, F., Ingram, I., Deane, F.P., Baker, A.L., & Osbourne, B.A. (2017). The Client Satisfaction Questionnaire-8: Psychometric properties in a cross-sectional survey of people attending residential substance abuse treatment. *Drug and Alcohol Review*, doi: 10.1111/dar.12522

Knabb, J., Welsh, R. & Graham-Howard, M. (2011). Treatment alternatives for mentally disordered offenders: a literature review. *Psychology, 2,* 122-131. doi:10.1002/cbm.594

Lancaster, G.A., Campbell, M.J., Eldridge, S., Farrin, A., Marchant, M., Muller, S., Perera, R., Peters, T.J., Prevost, A.T., & Rait, G. (2010). Trials in primary care: issues in the design of complex interventions. *Statistical Methods in Medical Research*, *19,* 349-77. doi: 10.1177/0962280209359883

Lawday, R., & Compton, S.J. (2013). Integrating models for integrated care pathways. In: S. Compton Dickinson, H. Odell-Miller and J. Adlam [Eds.] *Forensic music therapy: A treatment for men and women in secure hospital settings.* London, UK: Jessica Kingsley.

MacLehose, R.R. Reeves, B.C., Harvey, I.M., Sheldon, T.A., Russell, I.T., & Black, A.M. (2000). A systematic review of comparisons of effect sizes derived from randomised and non-randomised studies. *Health Technology Assessment, 4,* 1-154.

Meadows, J., & Kellett, S. (2017). Development and evaluation of cognitive analytic guided self-help (CAT-SH) for use in IAPT services. *Behavioural and Cognitive Psychotherapy, 45,* 266- 284. doi.org/10.1017/S1352465816000485

McGauley, G., & Bartlett, A. (2009). *Forensic mental health systems: concepts, systems and practice*. Oxford, UK: Oxford University Press.

Moher, D., Hopewell, S., Schulz, K.F., Montori, V., Gtzsche, P.C., Devereaux, P.J., Elbourne, D., Egger, M., & Altman, DG (2010). CONSORT 2010 explanation and

elaboration: updated guidelines for reporting parallel group randomised trials. *British Medical Journal, 340*. c869. doi: 10.1136/bmj.c869

Moffett, J. K., Torgerson, D., & Bell-Syer, S. (1999). Randomised controlled trial of exercise for low back pain: clinical outcomes, cost and preferences. *British Medical Journal*, *319:*279. doi: org/10.1136/bmj.319.7205.279

Ministry of Justice (2011). *Achieving best evidence in criminal proceedings: guidance on interviewing victims and witnesses, and guidance on using special measures* (accessed on 20/03/17 at

https://www.cps.gov.uk/publications/docs/best_evidence_in_criminal_proceedings.pdf

Mulder, S., & Chanen, A.M. (2013). Effectiveness of cognitive analytic therapy for personality disorders. *The British Journal of Psychiatry*, 202, 89-90. doi: 10.1192/bjp.bp.112.113571.

National Institute for Health and Care Excellence (2007). *Antisocial personality disorder scope consultation table*. Accessed 02/03/17 at https://www.nice.org.uk/ guidance/cg77/documents/scope-consultation-table- with-responses2

National Institute for Health and Care Excellence (2009). *Guidance 77: antisocial personality disorder: treatment management and prevention*. Accessed on 01/03/17 at http://www.rcpch.ac.uk/sites/default/ les/ asset_library/Research/Clinical%20Effectiveness/ CG077FullGuidelineX.pdf

Perepletchikova F, Treat, T.A., & Kazdin, A.E. (2007). Treatment integrity in psychotherapy research: analysis of the studies and examination of the associated factors. *Journal of Consulting and Clinical Psychology, 75*, 829-41. doi:10.1037/0022-006X.75.6.829

Pollock, P., & Belshaw, T. (1998). Cognitive analytic therapy for offenders. *The Journal of Forensic Psychiatry, 9*, 629-642. doi: org/10.1080/09585189808405378

Pollock, P., Broadbent, M., Clarke, S., Dorrian, A., & Ryle, A. (2001). The personality structure questionnaire (PSQ): A measure of the multiple self-states model of identity disturbance in cognitive analytic therapy. *Clinical Psychology and Psychotherapy, 8,* 59-72. doi: 10.1002/cpp.250

Pollock, P.H., Stowell-Smith, M., & Gopfert, M. (2006). *Cognitive analytic therapy for offenders: a new approach to forensic psychotherapy*. London, UK: Routledge.

Quimby, L.G., & Putnam, F.W. (1991). Dissociative symptoms and aggression in a state mental hospital. *Dissociation, 4*, 21–24.

Robb, S.L., Hanson-Abromeit, D., May, L. Hernandez-Ruiz, E., Allison, M., Beloat, A., Daugherty, S., Kurtz, R. Ott, A., OladimejiOyedele, O., Polasik, S., Rager, A., Rifkin, J., & Wolf, E. (2018). Reporting quality of music intervention research in healthcare: A systematic review. *Complementary Therapies in Medicine, 38*, 24-41. doi: org/10.1016/j.ctim.2018.02.008

Rosenblatt, A., & Attkisson, C.C. (1993). Assessing outcomes for sufferers of severe mental disorder: a conceptual framework and review. *Evaluation Programme Planning, 16*, 347-363. doi: org/10.1016/0149-7189(93)90049-E

Ryle, A. (1985). Cognitive theory, object relations and the self. *British Journal of Medical Psychology, 58*, 1-7. doi: 10.1111/j.2044-8341.1985.tb02608.x

Ryle, A. (1991). Object relations theory and activity theory: A proposed link by way of the procedural sequence model. *British Journal of Medical Psychology, 64*, 307-316. *doi/10.1111/j.2044-8341.1991.tb01669.x*

Ryle, A. (1997). *Cognitive analytic therapy and borderline personality disorder*. Chichester: John Wiley and Sons.

Ryle, A., & Kerr, I. (2002). *Introducing cognitive analytic therapy*. Chichester, UK: John Wiley and Sons.

Ryle, A. (2007). Investigating the phenomenology of borderline personality disorder with the states description procedure: clinical implications. *Clinical Psychology and Psychotherapy, 14*, 329-341. doi: 10.1002/cpp.54

Ryle, A. Kellett, S. Hepple, J., & Calvert, R. (2014). Cognitive analytic therapy at 30. *Advances in Psychiatric Treatment, 20,* 258-268. doi: 10.1192/apt.bp.113.011817

Ryle, A., & Kellett, S. (2018). Cognitive analytic therapy. In J.W. Livesley & R. Larstone (eds) *Handbook of personality disorders; theory, research and treatment (2nd edition).* Guilford Press.

Shepherd, G., Boardman, J. & Slade, M. (2008). *Making recovery a reality*. London: Sainsbury Centre for Mental Health.

Silverman, M. (2008). Quantitative comparison of cognitive behavioural therapy and music therapy research: a methodological best-practices analysis to guide future investigation for adult psychiatric patients. *Journal of Music Therapy, 4*, 457-506.

Silverman, M. (2015). *Music therapy in mental health for illness management and recovery.* Oxford University Press.

Talwar, N., Crawford, M., Maratos, A., Nur, U., McDermott, O., & Procter, S. (2006). Music therapy for in-patients with schizophrenia: exploratory randomised controlled trial. *British Journal of Psychiatry, 189*, 405–9. doi: 10.1192/bjp.bp.105.015073

Thabane, L., Ma, J., & Chu, R. (2010). A tutorial on pilot studies: the what, why and how. *BMC Medical Research Methodology*, *10*, 1. doi: 10.1186/1471-2288-10-1

Thomas, S.D.M. (2005). *Inverse care? Comparing needs and satisfaction with services between prisoners and patients in contact with forensic mental health services.* PhD. London, UK: Institute of Psychiatry, King's College

Torgerson, D., & Sibbald, B. (1998). Understanding controlled trials: what is a patient preference trial*? British Medical Journal*, *316*, 360. doi: org/10.1136/bmj.316.7128.360

Tseng, P-T., Chen,Y-W., Lin, P-Y., Tu, K-Y., Wang, H.Y., Cheng, Y-S., Chang, Y-C., Chang, C-H., Chung, W., & Wu, C.K. (2016). Significant treatment effect of adjunct music therapy to standard treatment on the positive, negative, and mood symptoms of schizophrenic patients: a meta-analysis. *BMC Psychiatry*, 16:16. doi: 10.1186/s12888-016-0718-8

Vernham, Z., Tapp, J., & Moore, E. (2015). Observer ratings of interpersonal behavior as predictors of aggression and self-harm in a high-security sample of male forensic inpatients. *Journal of Interpersonal Violence*, 31, 1597–1617. doi: org/10.1177/0886260515569060

Vygotsky, L. S. (1978). *Mind in society: the development of higher psychological processes.* Cambridge, Mass.: Harvard University Press.

Westbrook, D., & Kirk J. (2005). The clinical effectiveness of cognitive behaviour therapy: Outcome for a large sample of adults treated in routine practice. *Behaviour Research and Therapy, 43*, 1243-1261. doi: 10.1016/j.brat.2004.09.006